AP Statistics Practice Test (page 203)

T3.1 d. A correlation of near zero indicates no (or little) linear relationship, either positive or negative. Answers a, b and e indicate some form of linear relationship. Answer c implies no relationship whatsoever. It is possible to have a zero correlation where there is a strong relationship, just not a linear one.

T3.2 e. This point is influential because it is well above the mean for the amount spent on tobacco and well below the mean for the amount spent on alcohol. The observation (4.5, 6.0) is not an outlier because it does not have the greatest value in either dimension, nor does it fall outside the main pattern of the data set.

T3.3 c. This is the definition of r^2.

T3.4 a. The slope of the least-squares line depends on which variable is the explanatory variable and which is the response. Also, the slope $b = r\frac{s_y}{s_x}$ so $\frac{b}{r} = \frac{0.865}{0.79} = 1.09 = \frac{s_y}{s_x}$ which implies that $s_y > s_x$.

T3.5 a. The line predicts that a fish would have activity level $\hat{y} = 148.62 - 3.2167(20) = 84.286$. Looking at the residual plot, the fish with a predicted activity level of about 84 has a residual of approximately +3. Because $\text{residual} = y - \hat{y}$, $3 = y - 84$ and $y = 87$.

T3.6 c. This is another way of saying the typical error between the actual values and the predicted values using the linear model.

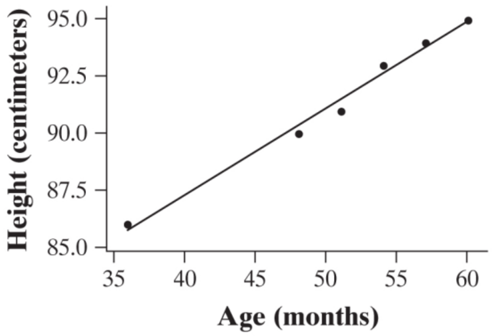
T3.7 b. The correlation does not have units attached to it.

T3.8 e. Because $b = r\frac{s_y}{s_x}$ and both standard deviations are 1, the correlation will be the same as the slope.

T3.9 b. The slope gives the predicted increase in y for each 1 unit increase in x. So if x increases 5 units, then the predicted value of y increases 5 times the slope or $5(3.3) = 16.5$ degrees F.

T3.10 c. Ethiopia is pulling the right hand side of the least-squares line up. If we remove that point, the line will be more flat, resulting in a decrease in the slope and an increase in the y intercept.

T3.11 (a) A scatterplot, with the regression line, is shown below.



(b) The regression line for predicting y = height from x = age is $\hat{y} = 71.95 + 0.3833x$.

(c) At age 480 months, we predict Sarah's height to be $\hat{y} = 71.95 + 0.3833(480) = 255.934\text{cm}$.

There are 2.54 cm to the inch, so her height in inches would be $\dfrac{255.934}{2.54} = 100.76\text{in}$.

(d) This height is impossibly large (about 8 feet, 5 inches) because it was an extrapolation. Obviously the linear trend will not continue until she is 40 years old. Our data was based only on the first 5 years of life and predictions should only be made for ages 0–5.

T3.12 (a) The unusual point is the one in the upper right hand corner with isotope value about -19.3 and silicon value about 345. This point is unusual in that it has such a high silicon value for the given isotope value.

(b) (i) If the point were removed the correlation would get closer to –1 because it does not follow the linear pattern of the other points. (ii) Because this point is "pulling up" the line on the right side of the plot, removing it will make the slope steeper (more negative) and make the y intercept smaller. Note that the y axis is to the *right* of the points in the scatterplot. (iii) Because this point has a large residual, removing it will make the size of the typical residual (s) a little smaller.

T3.13 (a) The equation of the least-squares regression line is $\hat{y} = 92.29 - 0.05762x$, where y represents the percent of the grass burned and x represents the number of wildebeest.

(b) The slope of the regression line says that for every increase of 1000 wildebeest, the predicted percent of grassy area burned decreases by about 0.058.

(c) Because the association is negative, the correlation is $r = -\sqrt{0.646} = -0.804$. There is a strong, negative linear association between the percent of grass burned and the number of wildebeest.

(d) Yes. Because there is no obvious leftover pattern in the residual plot, a linear model is appropriate for describing the relationship between wildebeest abundance and percent of grass area burned.